STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

23rd International Conference on Science and Technology Indicators (STI 2018)
"Science, Technology and Innovation indicators in transition"
12 - 14 September 2018 | Leiden, The Netherlands
#STI18LDN

Ethical challenges around the use of social media data: Views of researchers and research ethics committee members

Samuel, G.[1], Derrick, G.E.[2] and van Leeuwen, T.[3]

[1]gabrielle.samuel@kcl.ac.uk
Department of Global Health & Social Medicine, King's College London (United Kingdom)
[2]g.derrick@lancaster.ac.uk
Centre for Higher Education Research & Evaluation, Lancaster University (United Kingdom)
[3] leeuwen@cwts.leiden.nl
CWTS, Leiden University (The Netherlands)

Introduction

This paper reports the presence of a problematic "personal ethics" approach to decision-making by social media scholars who use data from general audience social media platforms for their research.

When new methodological tools like social media (SM) research are developed, differing norms of what constitutes ethically responsible research clash. Mining data from SM platforms such as Facebook, is of particular interest to scholars in the research evaluation field and the availability of this data has contributed to the rapid development of the field over recent years (Warren, Raison, & Dasgupta, 2017). More generally, SM data has been described as particularly rich sources of data (Vayena, Mastroianni, & Kahn, 2012), and the platforms used to generate it effective ways to recruit a large number of participants (Chu & Snider, 2013), as well as specifically sample participants with particular interests and/or health conditions (Renton et al., 2014; Rice et al., 2014)

Along with the growth of the field, there has been little focus on the ethical boundaries around using SM data. In the research evaluation field, there has been no discussion about the ethical considerations associated with using data from academic social media platforms. Instead, SM platforms are widely considered a general source of seemingly publicly available data (Gabaron, Serrano, Wynn, & Lau, 2014), and therefore free for public use. Recent behaviours of researchers and their affiliated entrepreneurial partners such as Cambridge Analytica

(Greenfield, 2018) however, highlight that despite dominant thought to the contrary, the use of SM data is not ethically-free, nor should it be free from ethical scrutiny.

Unfortunately, the nature of SM data is not contingent with traditional notions of consent, privacy, risk to participants, nor the freedom to withdraw. Nonetheless, there are now numerous calls for more researchers to have more consideration for the ethical issues associated with large-scale use of SM data. Indeed, the ethical hurdles related to SM research, these guidelines purport a researcher driven, case based approach.

This paper reports uses data from UK higher education institutions, journal Editors, UK funding bodies, social media research scholars, and research ethics committee members, consider the ethical boundaries of SM data. In particular, it focuses on how these academic governing bodies are failing to consistently promote a community-wide norm relating to the used of SM data. This leads to the application of ad hoc ethical definitions, or the promotion of the ability of researchers to apply a sense of "personal ethics" about SM use, that is not necessarily in line with the nature of SM data. As such, this paper offers a number of recommendations for the academic community to reflect on the ethical dimensions associated with responsible research behaviour relative to the use of SM data.

Methods

A mixed methods analysis combining guideline analysis, bibliometrics and qualitative (interview) techniques was employed in this study. Interviews from scholars using SM data for health research (n=14) were used along with interviews from members of university research ethics committees (n=19).

Document analysis of existing research ethics policies from Publishing houses, Peer-reviewed journal Editors and the UK funding councils were gathered. This included the top 20 UK research-intensive universities; the main UK funding bodies (UKRI, the Wellcome Trust and the National Institute of Health Research); the top 5 most prolific publishing houses, which account for more than 50% of all papers published; and the top 10 journals which publish SM research (n=16).

A summary of the data types, its nature and its source are included in Table 1.

Table 1: Sources and nature of data used in this study

Ethics ecosystem actor	Nature of methodological inquiry	Source
Research ethics committee	Interviews (qualitative)	REC members
Universities	Website analysis and clarifying survey	University ethics managers
Researchers	Interviews (qualitative)	Bibliometric search of authors of SM articles
Publishing houses	Document analysis and clarifying survey	Selective sample
Journals	Document analysis and clarifying survey	Bibliometric search for journals publishing SM articles
UK Funding councils	Document and website analysis	All RCUK actors

Analysis

Analysis of interview data was approached using inductive reasoning employing the inductive approach of grounded theory (Glaser & Strauss, 1967). The analysis (or coding) of data was based on two inter-linked rounds: overview analysis and detailed analysis (Glaser & Strauss, 1967). Overview analysis consisted of memo-making and broad coding. Extensive memo-making was employed by the interviewer directly after each interview. Broad coding proceeded by scanning the interview transcripts for relevant ideas and themes before each of these themes were explored in a detailed analysis.

Results

Researchers views

Researchers' experiences made them acutely aware of the various ethical issues associated with using SM approaches in their research. They were also familiar with the various guidelines published to help them negotiate these ethical issues, though in line with previous reports (Woodfield et al., 2013) these were generally viewed as not overly helpful. This was because the guidelines often acknowledged the messiness and complexity of SM research, described the ethical issues, but did little to prescribe how to act, or define boundaries relating to SM use. Researchers described them as *'deliberately grey' (researcher 1)*, feeling that they were left with little instruction on how to approach their research ethically: *'those ones (guidelines) and the BSA, I think those are the only two, They weren't very…they weren't kind of guidelines are they, they're more like make your own decision' (researcher 6)*

This lack of any prescriptive standard, alongside the need to adopt a researcher-led approach, gave interviewees permission to approach ethical decision-making in terms of how they felt ‘personally’ about their ethical choices, and the types of ethical issues which were salient to them within their own work:

> *‘…there’s a series of ethical dimensions and considerations and it is up to each researcher to reflect on that and consider what they mean for their particular project for their particular research question and population and specific methodology’ (researcher 3).*

Interviewees thus placed emphasis on the subjective, individual nature of ethics when justifying their research practices using SM data; *It’s my guidelines. Everybody has their own definition of ethics…. (researcher 5).* And interviewee 10 spoke about how any researcher, with any subjective beliefs about ethics, could justify their SM research just by ‘working around’ the vague guidelines:

> *‘they are [guidelines] slightly contradictory in places and you can argue around them...so there is some issues around integrity that don’t necessarily hold true and you can argue either way for some of the issues around that’.*

This meant that the ethical weight placed on any particular issue within a particular research project varied amongst researchers. A ‘personal ethics’ approach also led to inconsistencies in terms of whether researchers felt the need to receive ethical approval for SM research proposals[1]. For some, all research should be subject to ethical review *‘in some form’ (researcher 7),* since *‘it’s still research, so I would say anything that involves – even if it’s looking at newspaper coverage…should still be subject to ethical approval’ (researcher 2)*. Other interviewees did not feel the need to gain ethical approval for their research: *‘for data scraping I think no [I would not get ethical approval]’ (researcher 9).*

Research ethics committees

REC members were increasingly aware of the types of ethical issues associated with the use of SM data in health research, however they have little experience in reviewing SM research proposals. REC members spoke about the newness of the SM field, and the need to draw on

---

[1] Not all researchers’ beliefs about ethical review will affect their practices since some institutions demand ethical review for SM health research. However institutional practices are inconsistent and under-developed making it easier for researchers’ personal views about whether to receive ethical approval manifest in research practice

available guidelines, literature or training to aid with their decision-making. Though guidelines adopting a researcher-led approach were often described as being 'purposively vague' providing little direction in terms of decision-making.

Without useful guidelines, and parallel with the approaches adopted by a number of institutions and publishing houses discussed above, ethical decision-making was performed on a case-by-case basis *('it's very much on a case-by-case' (interviewee 11)).*

> *We are aware that this [their revised best practice guidance for research which uses social media data] was quite vague at the moment still. We tried to address the most significant issues that have come up, but because each case is different we can't say "well, you have to do this every single time" because it might not be appropriate depending on the subject area or the methodology or what people want to do (interviewee 17)*

Within this framework of ethical decision-making, focus was placed on researchers' justifications of their research approach, including their rationale for chosen research questions, methodologies and publication practices:

> *I think it depends on the project…you can't make blanket judgment and I think that's why we look at applications in detail in each case. And sometimes make different decisions even for projects that look pretty similar. It's how they build up their case doing that particular project (interviewee 15)*

Without any accepted guidance, and with inconsistencies in REC interviewees' approaches to ethical practice, researcher justifications of their research approach rather than objective standards, were often used to drive the process of ethical decision-making in the absence of firm guidelines on the ethical use of SM data. This allowed for researchers' 'personal ethics' approach to dominate the field's regulation of research behaviour around the use of SM data

Discussion

This paper shows how researchers and research governance bodies are currently not working effectively to collectively promote examples of responsible research behaviour around the use of SM data. Specifically, the tendency of social media research ethics guidance to promote a case-based, researcher-led ethical decision-making has led to the establishment of a 'personal ethics approach'

. This means that researchers generally do not submit their research proposals to formal governance structures and instead govern their own research behaviour that may not always be in line with a broader academic definition of ethically responsible behaviour. In addition, the inexperience of governing bodies in assessing whether research using SM data is being conducted ethically or not promotes this personal ethics.

This "personal ethics" approach as the main approach to regulating research using SM data can be dangerous as it leaves researchers at risk of failing to engage substantively about the ethical decisions they need to make when conducting research. To address this issue, the following guidelines are suggested for researchers, research organisations, funding organisations, journals and professional organisations, to address the risk;

1. Research using Social Media data should always seek formal, institutional ethical approval for their research prior to mining the data from a collaborative platform;
2. As with other big data resources such as biobanking, it must not be assumed that the notion of consent in research using SM data is the same as traditional definition of consent.
3. Ethical decisions concerning risk to participants, consent and privacy should focus beyond principles of how data is "collected" and instead concentrate of how data is being "used", processed and with whom it is shared.
4. Where traditional consent is not possible, other forms of consent and openness must be considered to allow participant to be aware of how their data is being used.
5. Researchers should not conduct themselves in one way as academics and another as entrepreneurs. In all situations, if a researcher is affiliated with a public funding body (such as a university), they are expected to conduct themselves in line with the highest academic standards of responsible research.

Although contentious, requiring all SM research to pass through a formalised process of ethical oversight will ensure that the next generation of knowledge, learning, and eventual standards relating to the use of SM data can emerge. While such a (temporary) change in governance may not solve all concerns related to how to conduct SM research ethically, it provides an important first step towards building an understanding within the academic community about the ethical boundaries surrounding the appropriate use of SM data.

References

Chu, J. L., & Snider, C. E. (2013). Use of a Social Networking Web Site for Recruiting Canadian Youth for Medical Research. Journal of Adolescent Health, 52, 792e794.

Gabaron, E., Serrano, J. A., Wynn, R., & Lau, A. (2014). Tweet Content Related to Sexually Transmitted Diseases: No Joking Matter. Journal of medical Internet research, 16(10), e228.

Glaser, B. G., & Strauss, A. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago: Aldine.

Greenfield, P. (2018, March 26th). The Cambridge Analytica files: the story so far. The Guardian. Retrieved from https://http://www.theguardian.com/news/2018/mar/26/the-cambridge-analytica-files-the-story-so-far

Renton, T., Tang, H., Ennis, N., Cusimano, M. D., Bhalerao, S., Schweizer, T. A., & Topolovec-Vranic, J. (2014). Web-Based Intervention Programs for Depression: A Scoping Review and Evaluation. Journal of medical Internet research, 16(9), e209. doi: v16i9e209 [pii]

Rice, S. M., Goodall, J., Hetrick, S. E., Parker, A. G., Gilbertson, T., Amminger, G. P., . . . Alvarez-Jimenez, M. (2014). Online and Social Networking Interventions for the Treatment of Depression in Young People: A Systematic Review. Journal of medical Internet research, 16(9), e206. doi: v16i9e206 [pii]

Vayena, E., Mastroianni, A., & Kahn, J. (2012). Ethical Issues in Health Research With Novel Online Sources. American Journal of Public Health, 102(12), 2225-2230.

Warren, H. R., Raison, N., & Dasgupta, P. (2017). The rise of altmetrics. JAMA, 317(2), 131-132.

Woodfield, K., Morrell, G., Metzler, K., Blank, G., Salmons, J., Finnegan, J., & Lucraft, M. (2013). Blurring the boundaries? New social media, new social research: Developing a network to explore the issues faced by researchers negotiating the new research landscape of online social media platforms. In N. C. f. R. M. M. R. paper (Ed.).